PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/36, C30B 19/00, 29/58, C07K 1/04, 17/02

A1

(11) International Publication Number: WO 96/17055

(43) International Publication Date: 6 June 1996 (06.06.96)

(21) International Application Number: PCT/US95/15714

(22) International Filing Date: 1 December 1995 (01.12.95)

(30) Priority Data:
08/349,636 2 December 1994 (02.12.94) US

(71) Applicant: VANDERBILT UNIVERSITY [US/US]; 405 Kirkland Hall, Nashville, TN 37240 (US).

(72) Inventors: HAWIGER, Jack, J.; 1818 Laurel Ridge Drive, Nashville, TN 37215 (US). DONAHUE, John, P.; 1601 Observatory Drive, Nashville, TN 37215 (US). PATEL, Hareshkumar; 2501 Soldiers Home Road #2D, West Lafayette, IN 47906 (US). ANDERSON, Wayne, F.; 1354 Burton Valley Road, Nashville, TN 37215 (US).

(74) Agents: NEEDLE, William, H. et al.; Needle & Rosenberg, Suite 1200, 127 Peachtree Street, N.E., Atlanta, GA 30303-1811 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: CARRIER PROTEIN DRIVEN CRYSTALLIZATION OF A PEPTIDE OR POLYPEPTIDE

(57) Abstract

The invention relates to methods of crystallizing a peptide or polypeptide comprising linking a peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein and crystallizing the chimeric protein. The invention also includes methods for determining the three-dimensional structure of a peptide or polypeptide, the crystallized chimeric proteins themselves, and methods for designing a peptide or polypeptide for screening for improved binding to a molecule by using the three dimensional structural information.

B3

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AT** | Austria | **GB** | United Kingdom | **MR** | Mauritania |
| **AU** | Australia | **GE** | Georgia | **MW** | Malawi |
| **BB** | Barbados | **GN** | Guinea | **NE** | Niger |
| **BE** | Belgium | **GR** | Greece | **NL** | Netherlands |
| **BF** | Burkina Faso | **HU** | Hungary | **NO** | Norway |
| **BG** | Bulgaria | **IE** | Ireland | **NZ** | New Zealand |
| **BJ** | Benin | **IT** | Italy | **PL** | Poland |
| **BR** | Brazil | **JP** | Japan | **PT** | Portugal |
| **BY** | Belarus | **KE** | Kenya | **RO** | Romania |
| **CA** | Canada | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CF** | Central African Republic | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CG** | Congo | | | **SE** | Sweden |
| **CH** | Switzerland | **KR** | Republic of Korea | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **KZ** | Kazakhstan | **SK** | Slovakia |
| **CM** | Cameroon | **LI** | Liechtenstein | **SN** | Senegal |
| **CN** | China | **LK** | Sri Lanka | **TD** | Chad |
| **CS** | Czechoslovakia | **LU** | Luxembourg | **TG** | Togo |
| **CZ** | Czech Republic | **LV** | Latvia | **TJ** | Tajikistan |
| **DE** | Germany | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **DK** | Denmark | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **US** | United States of America |
| **FI** | Finland | **ML** | Mali | **UZ** | Uzbekistan |
| **FR** | France | **MN** | Mongolia | **VN** | Viet Nam |
| **GA** | Gabon | | | | |

# CARRIER PROTEIN DRIVEN CRYSTALLIZATION OF A PEPTIDE OR POLYPEPTIDE

## BACKGROUND OF THE INVENTION

### FIELD OF THE INVENTION

This invention relates generally to crystallization of a peptide or polypeptide by use of a carrier protein. In one embodiment, the protein domain of human fibrinogen is crystallized using a carrier protein.

### BACKGROUND ART

Human fibrinogen ($M_r$ 340,000), a clottable protein in plasma and the most abundant ligand for the integrin $\alpha_{IIb}\beta_3$ (gpIIb/IIIa) receptor on platelets, is composed of pairs of three nonidentical polypeptide chains ($\alpha$, $\beta$, $\gamma$) that are extensively linked by disulfide bonds to form an elongated dimeric structure (for review see ref. 1). The binding of fibrinogen to the $\alpha_{IIb}\beta_3$ integrin receptor on activated platelets results in platelet aggregation *in vitro* and the formation of platelet-fibrin thrombi *in vivo* (for review see ref. 2). The segment of fibrinogen responsible for binding to the platelet $\alpha_{IIb}\beta_3$ integrin and aggregation of activated platelets has been mapped to the carboxyl terminus of the $\gamma$ chain and pinpointed to the continuous 12 amino acid sequence encompassing residues 400 to 411. This segment is both necessary and sufficient for optimal reactivity with platelet $\alpha_{IIb}\beta$ (3-6). With regard to this, it is noteworthy that mutation of both RGD cell adhesion motifs in the $\alpha$ chain of recombinant fibrinogen does not effect the ability of this molecule to mediate the aggregation of activated platelets (6). The fibrinogen $\gamma$ chain segment ($\gamma^{397-411}$) also serves as the ligand for the

clumping receptor on pathogenic staphylococci (7) and bears donor and acceptor sites for Factor XIIIa-catalyzed crosslinking of fibrin (8).

Two-dimensional NMR analysis of free fibrinogen $\gamma^{400-411}$ peptide in solution indicates the presence of a type II β turn spanning residues $Gln^{407}$ to $Asp^{410}$ (9). Human fibrinogen has not been crystallized in its native form. However, crystals can be obtained after limited cleavage of the native molecule with bacterial protease (10). Analysis of x-ray diffraction data generated using these crystals produced a model of the fibrinogen structure at 1.8 nanometer resolution. However, this x-ray diffraction analysis of crystals derived from proteolytically cleaved fibrinogen did not provide information about the three-dimensional structure of this biologically important segment (10). An attempt to determine the three dimensional structure of an RGD containing sequence by inserting 4-12 amino acid residues from the RGD region of human fibronectin into the loop of human lysozyme was unsuccessful (28).

Human fibrinogen is one of the many peptides or polypeptides that can not be adequately crystallized to provide suitable crystals for x-ray diffraction analysis because, for example, its three-dimensional structure is too bulky. In addition, functional segments of peptides or polypeptides, such as the functional segment of human fibrinogen discussed above, can not be adequately crystallized to provide suitable crystals for high resolution x-ray diffraction analysis.

The present invention solves these problems in the art by providing a method for crystallizing peptides and polypeptides, even more preferably small, functional protein segments, so that their three-dimensional structure can be solved using x-ray diffraction analysis. Thus, the invention provides a much needed means to characterize peptides or polypeptides so that the peptides or polypeptides can be optimized for a particular application or to rationally generate molecules which may bind the peptide or polypeptide or its native protein.

## SUMMARY OF THE INVENTION

In accordance with the purpose(s) of this invention, as embodied and broadly described herein, this invention, in one aspect, relates to a method of crystallizing a peptide or polypeptide comprising linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein and crystallizing the chimeric protein.

The invention further provides a method for determining the three dimensional structure of a peptide or polypeptide to a resolution of 1 nanometer or better comprising the steps of linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, crystallizing the chimeric protein, and then analyzing the chimeric protein to determine the three dimensional structure of the crystallized peptide or polypeptide.

In yet another aspect, the invention provides a crystallized chimeric protein, wherein the chimeric protein comprises a peptide or polypeptide linked to a terminus of a crystallizable carrier protein, wherein the peptide or polypeptide is crystallized such that, the three dimensional structure of the crystallized peptide or polypeptide can be determined to a resolution of 1 nanometer or better, and wherein the three dimensional structure of the peptide or polypeptide alone can not be determined to a resolution of 1 nanometer or better.

In yet another aspect, the invention provides a method for designing a first peptide or polypeptide for screening for improved binding to a molecule, comprising the steps of:

(1) evaluating the three dimensional structure of a second peptide or polypeptide, which can be the same or different than the first peptide or polypeptide, which was produced by:

a) linking the second peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, and

b) crystallizing the chimeric protein; and

(2) synthesizing a first peptide or polypeptide based on the crystal structure of the second peptide or polypeptide, wherein the first peptide or polypeptide can be screened for having improved binding to the molecule.

Additional advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and together with the description, serve to explain the principles of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the binding of the Met-lysozyme-fibrinogen $\gamma^{398-411}$ chimeric protein to platelet integrin $\alpha_{IIb}\beta_3$ using an ELISA system. The assay was done as described in example 1 and the average of triplicate determinations are shown. Met-lysozyme-fibrinogen $\gamma^{398-411}$ plus $\alpha_{IIb}\beta_3$ (□), Met-lysozyme-fibrinogen $\gamma^{398-411}$ minus $\alpha_{IIb}\beta_3$ (□), CEW lysozyme plus $\alpha_{IIb}\beta_3$ (◇).

FIG. 2 shows the stereo view of the $F_o$ - $F_c$ electron density for residues 134 to 144 ($\gamma^{401-411}$) of the Met-lysozyme-fibrinogen $\gamma^{398-411}$ chimeric protein. Before phases were calculated, residues 134 to 144 were removed from the model and a slow cooling simulated annealing (3000° K to 300° K) refinement was performed using X-PLOR (20). The map is contoured at 1.6 σ and a final model is superimposed.

FIG. 3 shows the stereo view of a ribbon diagram of the Met-lysozyme-fibrinogen $\gamma^{398-411}$ chimeric protein. The CEW lysozyme molecule is shown in green and the carboxyl terminal fibrinogen $\gamma^{398-411}$ segment (residues 131-144 of the chimeric protein) is shown in magenta.

FIG. 4 shows the intramolecular and intermolecular hydrogen bond interactions involving the fibrinogen $\gamma^{398-411}$ segment (residues 131-144) of the Met-lysozyme-fibrinogen $\gamma^{398-411}$ chimeric protein. Numbering of the CEW lysozyme residues is increased by one due to the presence of the Met residue on the amino-terminus of the protein. *A*, lysozyme portion of the chimeric molecule to which the fibrinogen $\gamma^{398-411}$ segment is attached; *B*, fibrinogen $\gamma^{398-400}$ segment of a symmetry related chimeric molecule; *C* and *D*, lysozyme portion of symmetry related chimeric molecules.

## DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention may be understood more readily by reference to the following detailed description of preferred embodiments of the invention and the Examples included therein and to the Figures and their previous and following description.

It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

The term "$M_r$" is a term familiar to one of ordinary skill in the art and means relative molecular mass or molecular weight.

In one aspect, the invention relates to a method of crystallizing a peptide or polypeptide comprising linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein and crystallizing the chimeric

protein.

The invention further provides a method for determining the three dimensional structure of a peptide or polypeptide to a resolution of 1 nanometer or better comprising the steps of linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, crystallizing the chimeric protein, and then analyzing the chimeric protein to determine the three dimensional structure of the crystallized peptide or polypeptide.

In yet another aspect, the invention provides a crystallized chimeric protein, wherein the chimeric protein comprises a peptide or polypeptide linked to a terminus of a crystallizable carrier protein, wherein the peptide or polypeptide is crystallized such that, the three dimensional structure of the crystallized peptide or polypeptide can be determined to a resolution of 1 nanometer or better, and wherein the three dimensional structure of the peptide or polypeptide alone can not be determined to a resolution of 1 nanometer or better.

In yet another aspect, the invention provides a method for designing a first peptide or polypeptide for screening for improved binding to a molecule, comprising the steps of:

(1) evaluating the three dimensional structure of a second peptide or polypeptide, which can be the same or different than the first peptide or polypeptide, which was produced by:

a) linking the second peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, and

b) crystallizing the chimeric protein; and

(2) synthesizing a first peptide or polypeptide based on the crystal structure of the second peptide or polypeptide, wherein the first peptide or polypeptide can be screened for having improved binding to the molecule.

As used herein, the term "linking" as used in, for example, the phrase "linking a peptide or polypeptide to a terminus of a crystallizable carrier protein" includes linking a peptide or polypeptide to a carrier protein by protein or peptide chemistry, wherein the C-terminus of either (a) the carrier protein or (b) the peptide or polypeptide is covalently attached to the N-terminus of (a) the peptide or polypeptide or (b) the carrier protein, respectively, via a peptide bond. The term "linking" additionally includes the construction through recombinant DNA technology of a nucleic acid which encodes a chimeric or hybrid protein comprising a peptide or polypeptide covalently bound to a carrier protein, such that expression of the nucleic acid produces the hybrid protein. which is in one embodiment set forth in the working example below.

The peptide or polypeptide is linked to only one terminal end of the carrier protein. It is not intended that the peptide or polypeptide be inserted into the interior of the carrier protein, that is, it is not inserted between the ends of the carrier protein. Thus, the peptide or polypeptide is not linked at both of its ends to the carrier protein or to segments of the carrier protein.

One method of constructing a chimeric protein for crystallization of an otherwise noncrystallizable binding domain or any other peptide or polypeptide is to synthesize a recombinant DNA molecule which encodes the chimeric protein. For example, oligonucleotide synthesis procedures are routine in the art and oligonucleotides coding for a particular protein region are readily obtainable through automated DNA synthesis. A nucleic acid for one strand of a double-stranded molecule can be synthesized and hybridized to its complementary strand. One can design these oligonucleotides such that the resulting double-stranded molecule has either internal restriction sites or appropriate 5' or 3' overhangs at the termini for cloning into an appropriate vector. Double-stranded molecules coding for relatively large proteins can readily be synthesized by first constructing several different double-stranded molecules that code for particular regions of the protein, followed by ligating these DNA molecules together. For example, *Cunningham, et al.*, "Receptor and Antibody Epitopes in Human Growth Hormone Identified by Homolog-Scanning Mutagenesis," *Science*, Vol. 243, pp.

1330-1336 (1989), have constructed a synthetic gene encoding the human growth hormone gene by first constructing overlapping and complementary synthetic oligonucleotides and ligating these fragments together. See also, *Ferretti, et al.*, Proc. Nat. Acad. Sci. 82:599-603 (1986), wherein synthesis of a 1057 base pair synthetic bovine rhodopsin gene from synthetic oligonucleotides is disclosed. Techniques such as this are routine in the art and are well documented. DNA fragments encoding peptides, polypeptides, or chimeric proteins can then be expressed *in vivo* or *in vitro* as discussed below.

An example of another method of obtaining a DNA molecule encoding a peptide or polypeptide for crystallization is to isolate that nucleic acid from the organism in which it is found. For example, a DNA or cDNA library can be constructed and screened for the presence of the nucleic acid of interest. Methods of constructing and screening such libraries are well known in the art and kits for performing the construction and screening steps are commercially available (for example, Stratagene Cloning Systems, La Jolla, CA). Once isolated, the nucleic acid can be directly cloned into an appropriate vector already containing the carrier protein. or if necessary, be modified to facilitate the subsequent cloning steps. Such modification steps are routine, an example of which is the addition of oligonucleotide linkers which contain restriction sites to the termini of the nucleic acid. General methods are set forth in Sumbrook *et al.*, "Molecular Cloning, a Laboratory Manual," Cold Spring Harbor Laboratory Press (1989). Alternatively, the peptide or polypeptide may be cloned into the appropriate vector prior to cloning of the carrier protein into that same vector. The carrier protein can either be chemically synthesized as described above, or isolated from an organism by the same procedures as described here for the peptide or polypeptide.

The peptide or polypeptide cloned into the vector can either encode native peptides or polypeptides, or as here, encode a chimeric protein where one region of the gene encode the carrier protein and another region of the gene encodes the peptide or polypeptide. For example, in one embodiment, the present invention comprises a chimeric protein which is composed of a protein region from chicken egg white

lysozyme linked to a binding domain region of the human fibrinogen γ gene. Other peptides or polypeptides can be linked to this region of chicken egg white lysozyme, or another carrier protein or protein region in a similar manner.

Once a nucleic acid encoding a particular peptide or polypeptide of interest, or a region of that nucleic acid, is constructed or isolated, that nucleic acid can then be cloned into an appropriate vector adjacent to the carrier protein, which can direct the *in vivo* or *in vitro* synthesis of that chimeric protein. The vector is contemplated to have the necessary functional elements that direct and regulate transcription of the inserted gene, or hybrid gene. These functional elements include, but are not limited to, a promoter, regions upstream or downstream of the promoter, such as enhancers that may regulate the transcriptional activity of the promoter, an origin of replication, appropriate restriction sites to facilitate cloning of inserts adjacent to the promoter, antibiotic resistance genes or other markers which can serve to select for cells containing the vector or the vector containing the insert, RNA splice junctions, a transcription termination region, or any other region which may serve to facilitate the expression of the inserted gene or hybrid gene. (See generally, Sambrook *et al.*)

There are numerous *E. coli* (Escherichia coli) expression vectors known to one of ordinary skill in the art which are useful for the expression of the nucleic acid insert. Other microbial hosts suitable for use include bacilli, such as *Bacillus subtilis*, and other enterobacteriaceae, such as *Salmonella*, *Serratia*, and various *Pseudomonas* species. In these prokaryotic hosts one can also make expression vectors, which will typically contain expression control sequences compatible with the host cell (e.g., an origin of replication). In addition, any number of a variety of well-known promoters will be present, such as the lactose promoter system, a tryptophan (Trp) promoter system, a beta-lactamase promoter system, or a promoter system from phage lambda. The promoters will typically control expression, optionally with an operator sequence, and have ribosome binding site sequences for example, for initiating and completing transcription and translation. If necessary, an amino terminal methionine can be provided by insertion of a Met codon 5' and in-frame with the downstream nucleic acid

insert. Also, the carboxy-terminal extension of the nucleic acid insert can be removed using standard oligonucleotide mutagenesis procedures.

Additionally, yeast expression can be used. There are several advantages to yeast expression systems. First, evidence exists that proteins produced in a yeast secretion systems exhibit correct disulfide pairing. Second, post-translational glycosylation is efficiently carried out by yeast secretory systems. The *Saccharomyces cerevisiae* pre-pro-alpha-factor leader region (encoded by the *MF"-1* gene) is routinely used to direct protein secretion from yeast (Brake *et al.*, "∝-Factor-Directed Synthesis and Secretion of Mature Foreign Proteins in *Saccharomyces cerevisiae*," PNAS, Vol. 81, pp. 4642-4646 (1984)). The leader region of pre-pro-alpha-factor contains a signal peptide and a pro-segment which includes a recognition sequence for a yeast protease encoded by the *KEX2* gene: this enzyme cleaves the precursor protein on the carboxyl side of a Lys-Arg dipeptide cleavage signal sequence. The nucleic acid coding sequence can be fused in-frame to the pre-pro-alpha-factor leader region. This construct is then put under the control of a strong transcription promoter, such as the alcohol dehydrogenase I promoter or a glycolytic promoter. The nucleic acid coding sequence is followed by a translation termination codon which is followed by transcription termination signals. Alternatively, the nucleic acid coding sequences can be fused to a second protein coding sequence, such as Sj26 or β-galactosidase, used to facilitate purification of the fusion protein by affinity chromatography. The insertion of protease cleavage sites to separate the components of the fusion protein is applicable to constructs used for expression in yeast. Efficient post translational glycosolation and expression of recombinant proteins can also be achieved in Baculovirus systems.

Mammalian cells permit the expression of proteins in an environment that favors important post-translational modifications such as folding and cysteine pairing, addition of complex carbohydrate structures, and secretion of active protein. Vectors useful for the expression of active proteins in mammalian cells are characterized by insertion of the protein coding sequence between a strong viral promoter and a polyadenylation signal. The vectors can contain genes conferring hygromycin resistance,

gentamicin resistance, or other genes or phenotypes suitable for use as selectable markers, or methotrexate resistance for gene amplification. The chimeric protein coding sequence can be introduced into a Chinese hamster ovary (CHO) cell line using a methotrexate resistance-encoding vector, or other cell lines using suitable selection markers. Presence of the vector DNA in transformed cells can be confirmed by Southern blot analysis. Production of RNA corresponding to the insert coding sequence can be confirmed by Northern blot analysis. A number of other suitable host cell lines capable of secreting intact human proteins have been developed in the art, and include the CHO cell lines, HeLa cells, myeloma cell lines, Jurkat cells, etc. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter, an enhancer, and necessary information processing sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional terminator sequences. Preferred expression control sequences are promoters derived from immunoglobulin genes, SV40, Adenovirus, Bovine Papilloma Virus, etc. The vectors containing the nucleic acid segments of interest can be transferred into the host cell by well-known methods, which vary depending on the type of cellular host. For example, calcium chloride transformation is commonly utilized for prokaryotic cells, whereas calcium phosphate, DEAE dextran, or lipofectin mediated transfection or electroporation may be used for other cellular hosts.

Alternative vectors for the expression of genes in mammalian cells, those similar to those developed for the expression of human gamma-interferon, tissue plasminogen activator, clotting Factor VIII, hepatitis B virus surface antigen, protease Nexinl, and eosinophil major basic protein, can be employed. Further, the vector can include CMV promoter sequences and a polyadenylation signal available for expression of inserted nucleic acids in mammalian cells (such as COS-7).

Expression of the gene or hybrid gene can be by either *in vivo* or *in vitro*. *In vivo* synthesis comprises transforming prokaryotic or eukaryotic cells that can serve as host cells for the vector. One example of a hybrid protein inserted into a prokaryotic *in vivo* expression vector is given in Example 1. The expression of the inserted gene is

regulated by the plasmid promoter activity, which itself is controlled by an inhibitor produced by the host cell. This inhibitor prevents expression of the gene by down-regulating the promoter, but by the addition of an inducer, here, isopropylthiogalactoside (IPTG), this negative regulation is overcome and the promoter initiates transcription. In this manner, one can control the temporal expression of the gene and therefore maximize the level of protein produced by the cells.

Alternatively, expression of the gene can occur in an *in vitro* expression system. For example, *in vitro* transcription systems are commercially available which are routinely used to synthesize relatively large amounts of mRNA. In such *in vitro* transcription systems, the nucleic acid encoding the hybrid protein would be cloned into an expression vector adjacent to a transcription promoter. For example, the Bluescript II cloning and expression vectors contain multiple cloning sites which are flanked by strong prokaryotic transcription promoters. (Stratagene Cloning Systems, La Jolla, CA). Kits are available which contain all the necessary reagents for *in vitro* synthesis of an RNA from a DNA template such as the Bluescript vectors. (Stratagene Cloning Systems, La Jolla, CA). RNA produced *in vitro* by a system such as this can then be translated *in vitro* to produce the desired polypeptide. (Stratagene Cloning Systems, La Jolla, CA).

Another method of producing a chimeric protein comprising a peptide or polypeptide linked to a carrier protein is to link two peptides or polypeptides together by protein chemistry techniques. For example, peptides or polypeptides can be chemically synthesized using currently available laboratory equipment using either Fmoc (9-fluorenylmethyloxycarbonyl) or Boc (*tert* -butyloxycarbonoyl) chemistry. (Applied Biosystems, Inc., Foster City, CA). One skilled in the art can readily appreciate that a peptide or polypeptide corresponding to a protein domain can be synthesized and then coupled to a peptide or polypeptide corresponding to a carrier protein by standard chemical reactions. For example, a peptide or polypeptide can be synthesized and not cleaved from its synthesis resin whereas the other fragment of a hybrid peptide can be synthesized and subsequently cleaved from the resin, thereby exposing a terminal group which is functionally blocked on the other fragment. By peptide condensation reactions,

these two fragments can be covalently joined via a peptide bond at their carboxyl and amino termini, respectively, to form a hybrid peptide. (Grant, G.A., "Synthetic Peptides: A User Guide," W.H. Freeman and Co., N.Y. (1992) and Bodansky, M. and Trost, B., Ed., "Principles of Peptide Synthesis," Springer-Verlag Inc., N.Y. (1993)). Alternatively, the peptide or polypeptide as well as the carrier protein can by independently synthesized *in vivo* as described above. Once isolated, these independent peptides or polypeptides and carrier proteins may be linked to form a chimeric protein via similar peptide condensation reactions.

For example, enzymatic ligation of cloned or synthetic peptide segments can allow relatively short peptide fragments to be joined to produce larger peptide fragments, polypeptides or whole protein domains (Abrahmsen, L., *et al.*, *Biochemistry*, Vol. 30, p. 4151 (1991)). Alternatively, native chemical ligation of synthetic peptides can be utilized to synthetically construct large peptides or polypeptides from shorter peptide fragments. This method consists of a two step chemical reaction (Dawson, *et al.*, "Synthesis of Proteins by Native Chemical Ligation," *Science*, Vol. 266, pp. 776-779 (1994)). The first step is the chemoselective reaction of an unprotected synthetic peptide-α-thioester with another unprotected peptide segment containing an amino-terminal Cys residue to give a thioester-linked intermediate as the initial covalent product. Without a change in the reaction conditions, this intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site. Application of this native chemical ligation method to the total synthesis of a protein molecule is illustrated by the preparation of human interleukin 8 (IL8) (Clark-Lewis, I., *et al.*, *FEBS Lett.*, Vol. 307, p. 97 (1987), Clark-Lewis, I., *et al.*, *J.Biol.Chem.*, Vol. 269, p. 16075 (1994), Clark-Lewis, I., *et al.*, *Biochemistry*, Vol. 30, p. 3128 (1991), and Rajarathnam, K., *et al.*, *Biochemistry*, Vol. 29, p. 1689 (1994)).

Alternatively, unprotected peptide segments can be chemically linked where the bond formed between the peptide segments as a result of the chemical ligation is an unnatural (non-peptide) bond (Schnolzer, M., *et al.*, *Science*, Vol. 256, p. 221 (1992)). This technique has been used to synthesize analogs of protein domains as well

as large amounts of relatively pure proteins with full biological activity (deLisle Milton, R.C., *et al.*, "Techniques in Protein Chemistry IV," Academic Press, New York, pp. 257-267 (1992)).

Any carrier protein that can be readily crystallized with the peptide or polypeptide can be used in this invention. Typically, the carrier protein is crystallizable alone or by itself (i.e., not linked to a peptide or polypeptide), and this crystallization property helps to drive the crystallization of the chimeric protein. The resultant chimeric protein is adequately crystallizable. Specifically, when the chimeric protein is crystallized, a very complete crystal lattice results, so that, the three dimensional crystal structure of the peptide or polypeptide can be determined at high resolution. The carrier protein is selected so that preferably *(i)* it is easily crystallized and its three-dimensional structure is well characterized, *(ii)* it can be expressed at high levels and is easily purified, *(iii)* it has amino or carboxyl termini that are solvent accessible so the chimeric protein will not require structural alterations, and *(iv)* it will readily form well ordered crystals with high solvent content.

The carrier protein can be of any size with no particular upper or lower limit as long as it can be readily crystallized with the peptide or polypeptide of this invention. The carrier protein is in one embodiment at least 10,000 $M_r$ (about 83 amino acids), in another embodiment from about 10,000 $M_r$ (about 83 amino acids) to about 250,000 $M_r$ (about 2,083 amino acids), in another embodiment from about 10,000 $M_r$ (about 83 amino acids) to about 135,000 $M_r$ (about 1,125 amino acids).

Typical carrier proteins that can be used or tested for use in the invention are chicken egg white lysozyme, human lysozyme, thioredoxin, glutathione-S-transferase, maltose binding protein, Fab antibody fragment, glycogen phosphorylase, purine nucleotide phosphorylase, beta galactosidase, or a derivative thereof. In a specific embodiment, the carrier protein is chicken egg white lysozyme (CEW) or Met-CEW.

The method of this invention is applicable to many biological structures. Typical structures that can be crystallized using the present method include peptides and polypeptides. As used herein, "peptide or polypeptide" is intended to include peptides, polypeptides, proteins, protein segments, protein domains, and the like. Additionally, as used herein, "peptide" and "polypeptide" includes mimetics. Some examples of peptides or polypeptides that can be used in this invention include, but are not limited to, Integrin $\alpha_{IIb}$ subunit, Integrin $\beta_3$ subunit, Integrin $\alpha_5$ subunit, Integrin $\beta_1$ subunit, Integrin $\alpha_L$ subunit, Integrin $\alpha_M$ subunit, Integrin $\beta_2$ subunit, CD14 (membrane protein binding lipopolysaccharide and other bacterial envelope constituents), Von Willebrand Factor segments interacting with platelet membrane receptors (integrins $\alpha_{IIb}\beta_3$ and non-integrin glycoprotein Ib/Ix complex), Fibronectins, Transcriptional factor NF-κB (IκBα), and inhibitors of Transcriptional factor NF-κB (IκBα). In one embodiment of the present invention, the peptide or polypeptide is the peptide set forth in SEQ ID NO: 1.

The peptide or polypeptide is not restricted to a particular size. In one embodiment, the peptide or polypeptide is at least 360 $M_r$ (about 3 amino acids), in another embodiment the peptide or polypeptide is 360 $M_r$ (about 3 amino acids) to 72,000 $M_r$ (about 600 amino acids), in another embodiment from 2,400 $M_r$ (about 20 amino acids to 36,000 $M_r$ (about 300 amino acids), in another embodiment 6,000 $M_r$ (about 50 amino acids) to 24,000 $M_r$ (200 amino acids).

The carrier protein is typically larger than the peptide or polypeptide. However, it is possible for the carrier protein to be the same size or smaller than the peptide or polypeptide if the combined carrier protein/peptide or polypeptide chimeric protein is of an adequate size to be crystallized.

Once the chimeric protein is formed, any technique used in the art to crystallize a peptide or polypeptide can be similarly used in the method of this invention to crystallize the chimeric protein. Typically, crystallization involves lowering the temperature of the chimeric protein until the chimeric protein forms a crystal lattice as provided in the Examples. General crystallization conditions are set forth in

*Crystallization of Nucleic Acids and Proteins: A Practical Approach*, ed. Ducruix *et al.*, IRL Press (1992).

Once the chimeric protein has been crystallized, any technique typically used in the art can be used to help discern the structure of the crystallized chimeric protein. Generally, crystallography techniques are employed, such as x-ray diffraction. Such techniques are generally set forth in Drenth, J., *Principles of Protein Xray Crystallography*, Springer Verlag (1994).

The present invention provides for the formation of highly organized crystal structures that are tightly packed, whereby the three dimensional structure of the chimeric protein and the peptide or polypeptide can be adequately determined, more particularly, the three dimensional structure can be determined to a high resolution. In particular, the three dimensional structure of the crystallized chimeric protein or the crystallized peptide or polypeptide portion of the crystallized chimeric protein is preferably determined to a resolution of 1.5 nanometers or better, more preferably 1 nanometer or better, even more preferably 0.5 nanometers or better, even more preferably about 0.24 nanometers or better. This invention is particularly applicable to peptides and polypeptides that can not be adequately crystallized alone. That is, without the use of a carrier protein, the peptides and polypeptides of this invention can not be adequately crystallized to determine their three dimensional structure. Alone, their crystal structure cannot be determined to a resolution of 1.5 nanometers or better, more preferably 1 nanometer or better, even more preferably 0.5 nanometers or better, even more preferably about 0.24 nanometers or better.

Utility:

The present invention has many useful and practical utilities as noted herein. For example, the present invention provides for adequately crystallizing a peptide or polypeptide so that its three dimensional structure can then be determined. This allows for the design of a peptide or polypeptide for screening for improved binding

to a molecule. The molecule can be, for example, a receptor. The binding can be used either to antagonize or agonize the receptor.

In a specific embodiment, the present invention provides for a method for crystallizing a functional protein segment of the human fibrinogen chain at a resolution of 0.24 nanometers. This segment encompasses the recognition site for the integrin $\alpha_{IIb}\beta_3$ receptor on activated platelets and for the clumping receptor on pathogenic staphylococci, and also bears donor and acceptor sites for Factor XIIIa-catalyzed crosslinking of fibrin. The structural information derived from this analysis provides a rational basis for the design of inhibitors of these important functions of fibrinogen. In general, carrier protein driven crystallization facilitates the determination of the three-dimensional structure of peptides and polypeptides, especially functional segments of other proteins, that are, like fibrinogen, difficult to adequately crystallize by other techniques.

Experimental

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the invention claimed herein are made and evaluated, and are intended to be purely exemplary of the invention and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.) but some errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C, and pressure is at or near atmospheric.

## Example 1

### Construction of the Met-Lysozyme-Fibrinogen $\gamma^{398-411}$ Expression

**Plasmid.** An expression plasmid was constructed in which a DNA sequence encoding the 14 carboxyl terminal residues of the fibrinogen γ chain was inserted downstream of

the CEW lysozyme cDNA. Lysozyme cDNA, derived from plasmid pls-1 (11) (provided by Günther Schütz, Max-Planck-Institut für Molekulare Genetik), was cloned into the prokaryotic expression vector pKP1500 (12) (provided by Takeyoshi Miki, Kyushu University) essentially as described (12,13). To facilitate the insertion of a DNA fragment encoding the carboxyl terminal amino acid residues of the fibrinogen γ chain, a *Pst*I site was inserted at the 3' end of the lysozyme coding sequence by oligonucleotide directed mutagenesis using M13 vectors (14). This insertion resulted in the addition of a glutamine to the carboxyl terminus of lysozyme. This residue was equivalent to $Gln^{398}$ of the fibrinogen γ chain (15) in the final construct. The 575 bp *Eco*RI/*Hin*dIII fragment from M13mp11 that contained the modified lysozyme cDNA was ligated into pKP1500. The resulting plasmid was named pNED6.

Based on the fibrinogen γ chain cDNA sequence (15), the following complementary oligonucleotides were synthesized, which are also set forth in Seq. ID Nos. 2 and 3.

```
    G CAA CAC CAC CTA GGG GGA GCC AAA CAG GCT GGA GAC GTT TA
AC GTC GTT GTG GTG GAT CCC CCT CGG TTT GTC CGA CCT CTG CAA ATT CGA
   gln gln his his leu gly gly ala lys gln ala gly asp val ***
```

After annealing, the resulting double stranded oligonucleotide contained a translation termination codon immediately following $Val^{411}$ and *Pst*I and *Hin*dIII sticky ends at 5' and 3' termini. This DNA fragment was ligated into *Pst*I/*Hin*dIII digested pNED6 to form plasmid pNED7. The DNA sequence of the lysozyme cDNA with the inserted γ chain oligonucleotide was confirmed using the dideoxy chain termination method (16).

**Met-Lysozyme-Fibrinogen $\gamma^{398-411}$ Expression and Purification.** *Escherichia coli* strain KP3998 (12) containing plasmid pNED7 was grown at 40°C in TYG broth (1% tryptone, 0.5% yeast extract, 0.25% glycerol, 1 mM $MgSO_4$, 0.1 M potassium phosphate, pH 7) plus 100 μg/ml ampicillin. Chimeric protein synthesis was induced by addition of IPTG (0.5 mM final concentration). Met-lysozyme-fibrinogen $\gamma^{398-411}$ was purified from cell pellets as described (13) except that an ultrafiltration step

was introduced after the third acetic acid extraction of the reduced protein. The usual yield of Met-lysozyme-fibrinogen $\gamma^{398-411}$ was 1.5 mg/g wet weight of cells with an estimated purity of at least 98%, as determined by Coomassie blue staining of these protein preparations in SDS-polyacrylamide gels.

**Met-Lysozyme-Fibrinogen $\gamma^{398-411}$ Renaturation.** Purified, reduced protein was renatured as described (13). The progress of the renaturation reaction was monitored by measuring the reconstitution of lysozyme enzymatic activity using a *Micrococcus lysodeikticus* lysis assay performed as described by Sigma Chemical Co. for CEW lysozyme. Maximal enzymatic activity was recovered after 1 hour of incubation. The renaturation reaction mixture was dialyzed against 0.1 M acetic acid and lyophilized. The residue was dissolved in 1 ml of 0.1 M acetic acid and chromatographed on a 1.5 X 47 cm column of Sephadex G-25 equilibrated and run with 0.1 M acetic acid. The Met-lysozyme-fibrinogen $\gamma^{398-411}$ which was contained in the excluded volume was lyophilized and then dissolved in 50 mM NaCl. The protein concentration was determined using an $A_{280}$ for a 1% solution of native CEW lysozyme of 26.3 (17).

**Measurement of Binding of Met-Lysozyme-Fibrinogen $\gamma^{398-411}$ to Platelet Integrin $\alpha_{IIb}\beta_3$.** Purified platelet fibrinogen receptor (integrin $\alpha_{IIb}\beta_3$, 10 $\mu$g/ml) (provided by David Phillips of COR Therapeutics, Inc.) in coating buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 0.05% $NaN_3$) was applied to the wells of Immulon 2 (Dynatech Laboratories, Inc.) microtiter plates. The plates were incubated at room temperature for 18 hours and then blocked with gelatin (20 $\mu$g/ml in coating buffer). Protein-coated wells were washed with TBSCT (20 mM Tris-HCl, pH 7.6, 137 mM NaCl, 0.5 mM $CaCl_2$, 0.05% Tween 20) and Met-lysozyme-fibrinogen $\gamma^{398-411}$ or CEW lysozyme were added. Plates were incubated for 2 hours at room temperature and washed with TBSCT. Chimeric protein or CEW lysozyme binding was detected using monoclonal anti-CEW lysozyme HyHel-5 (approximately 1.5 $\mu$g/ml) (kindly provided by Sandra Smith-Gill, National Cancer Institute) and goat anti-mouse IgG conjugated to alkaline phosphatase (A-3688, Sigma Chemical Co.). Secondary antibody binding was detected using *p*-nitrophenyl phosphate (1 mg/ml) dissolved in 1 M Tris-HCl (pH 9.5).

**Crystallization, Data Collection, and Structure Determination.** The purified, renatured Met-lysozyme-fibrinogen $\gamma^{398-411}$ (5 to 10 mg/ml, pH 2.5) was crystallized at 22°C in hanging drops using 1.4 M $(NH_4)_2SO_4$ as precipitant buffered with 0.1 M Tris-HCl (pH 8). Any increase in the ionic strength or pH of the chimeric protein preparation present in hanging drops resulted in rapid protein precipitation and the formation of extremely thin needle crystals. Although the hanging drop was equilibrated against ammonium sulfate buffered by 0.1 M Tris-HCl (pH 8), the exact pH of the drop when the one crystal large enough for data collection initially grew was not determined.

Data were collected on a single crystal (1.5 x 0.05 x 0.05 mm) at room temperature using a pair of San Diego Multiwire Systems area detectors. The crystal space group was $P2_12_12_1$ and the unit cell dimensions were a=55.9 Å, b=74.0 Å, and c=30.8 Å. The 2.4 Å diffraction data were collected with an R merge of 0.140 for 15,024 observations of 4625 unique reflections.

The structure was determined using the molecular replacement method (18) with CEW lysozyme coordinates (19) and molecular replacement routines in the X-PLOR (20) program system. An initial rotation search gave a peak of 5.4 $\sigma$ that was used for Patterson correlation refinement, which brought the correlation coefficient to 0.204. The refined rotation parameters were used in a translation search that yielded a peak of 10 $\sigma$ with an R value of 0.378.

Coordinates for the lysozyme portion of the structure, determined using molecular replacement, were subjected to an initial rigid body refinement. After rigid body refinement, Powell energy minimization converged at an R value of 0.242. Further refinement was carried out using a slow cooling simulated annealing procedure (3000°K to 300° K) (21) and B factor refinement resulting in a final R value of 0.202.

**Results:**

**Characterization of Met-Lysozyme-Fibrinogen $\gamma^{398-411}$ Expressed in *Escherichia coli*.** In an effort to determine the three-dimensional structure of the fibrinogen γ chain receptor recognition segment, a plasmid was constructed that directs the synthesis in *E. coli* of a chimeric protein composed of CEW lysozyme with a Met residue added to the amino-terminus and fibrinogen γ chain residues 398-411 added to the carboxyl terminus. Met-lysozyme-fibrinogen $\gamma^{398-411}$ was purified from cytoplasmic precipitates (12) in a fully reduced, denatured form and subsequently renatured by sulfhydryl-disulfide exchange *in vitro*. The specific activity of the renatured Met-lysozyme-fibrinogen $\gamma^{398-411}$ was 30% of renatured, native CEW lysozyme. This result was consistent with a previous report of Met-lysozyme renaturation *in vitro* (13).

It was important to establish that the γ chain sequence present as a carboxyl terminal extension of CEW lysozyme could adopt a biologically active conformation as measured by $\alpha_{IIb}\beta_3$ receptor binding. The low solubility of the chimeric protein at neutral pH prevented using an assay in which inhibition of $^{125}$I-labeled fibrinogen binding to $\alpha_{IIb}\beta_3$ on activated platelets was measured. However, using an ELISA system, it was demonstrated that binding of Met-lysozyme-fibrinogen $\gamma^{398-411}$ to immobilized $\alpha_{IIb}\beta_3$ was approximately 10-fold greater than the binding of native CEW lysozyme (Fig. 1). This binding was dependent on the presence of the $\alpha_{IIb}\beta_3$ receptor and on the concentration of Met-lysozyme-fibrinogen $\gamma^{398-411}$, reaching saturation between 40 and 80 μg/ml of protein.

**Determination of the Structure of Met-Lysozyme-Fibrinogen $\gamma^{398-411}$.** The structure of the chimeric protein was determined by molecular replacement methods (18) using the coordinates from the tetragonal CEW lysozyme structure (19). Comparison of the packing of the chimeric protein with that of CEW lysozyme in the orthorhombic crystal form (22) revealed that the chimeric molecules had rotated so that the carboxyl terminus entered a space between lysozyme molecules rather than abutting another molecule. The addition of the 14 carboxyl terminal residues of the fibrinogen γ

chain to the carboxyl terminus of the CEW lysozyme did not cause any significant changes in the structure of the lysozyme portion of the chimeric protein. The rms difference in the polypeptide backbone coordinates between the tetragonal lysozyme structure and the chimeric protein was 0.63 Å. The average B values for the lysozyme main chain and side chain atoms were 14.0 and 14.4, respectively; the average B values for the fibrinogen $\gamma^{398-411}$ segment (residues 131-144 of the chimeric protein) main chain and side chain atoms were 34.7 and 32.1, respectively. The higher temperature factors for the fibrinogen $\gamma$ chain segment of the chimeric molecule suggest that it exhibits greater mobility than the lysozyme to which it is attached. This is not surprising given the carboxyl terminal location of this segment and its relatively high glycine content.

Several approaches were taken to verify the validity of the refined model of the fibrinogen $\gamma^{398-411}$ segment. The real space fit correlation coefficient (23) for the $\gamma^{398-411}$ segment to the $2F_o - F_c$ electron density map ranged from 65 to 79%. The same calculation for the lysozyme portion of the structure gave values from 61 to 88%. This indicated that the fibrinogen $\gamma^{398-411}$ segment does not have a dramatically worse fit to the electron density than the lysozyme portion, which was used as the model to determine the crystal structure.

Cross validation of the structure using the reciprocal space Free R value (20,24) was also used to determine that the addition of the fibrinogen $\gamma^{398-411}$ segment to the model truly improved the agreement with the observed data rather than just adding more parameters. A randomly chosen 10% of the data were used for the calculation of the Free R value. For the lysozyme model itself, the Free R for residues 2-130 was 0.370 and for the final structure, residues 1-144, it was 0.328. To determine if most of the structure could be correctly placed and residues 136-144 ($\gamma^{403-411}$) incorrectly placed, a calculation of the Free R was done for a model with residues 2-135 correctly placed and an incorrect placement of residues 136-144. In this case the Free R was 0.392, while the Free R for residues 1-135 by themselves was 0.359. These results indicate that addition of the fibrinogen $\gamma^{398-411}$ segment (residues 131-144) to the model increased the agreement with the observed diffraction data.

A slow cooling simulated annealing omit map ($F_o$ - $F_c$) (25) was calculated using a model with residues 134-144 ($\gamma^{401-411}$) omitted. This map and the final model of the fibrinogen $\gamma^{398-411}$ segment are shown in Fig. 2. This again demonstrated that the model agrees with experimental data. One final check was that the stereochemistry of the model was acceptable. None of the fibrinogen γ chain residues (131-144) fall within disallowed regions of a Ramachandran plot (26,27).

**Structure of the Fibrinogen $\gamma^{398-411}$ Receptor Recognition Segment.**

The model of the carboxyl terminal extension of the chimeric protein was built into $2F_o$ - $F_c$ electron density maps in two stages. The first map was calculated using phases derived from the positioned CEW lysozyme model and allowed the addition of residues 131-134 ($\gamma^{398-401}$). This extended model was used to calculate phases and an electron density map like that shown in Fig. 2. This second map was good enough to place the rest of the carboxyl terminal residues. The general conformation of the fibrinogen $\gamma^{398-411}$ segment is a wide turn followed by an extended region and ending with a wide turn (Figs. 2 and 3). In each of these turns the first two residues are analogous to the first two residues of a β turn and the last two residues are analogous to the last two residues of a β turn. However, in these turns an additional residue separates the two halves, leaving them too far apart for hydrogen bond interactions to occur. For residues 131-135 ($\gamma^{398-402}$), which comprise the first turn, the $Gln^{132}$ ($\gamma^{399}$) side chain projects into the space between the two halves and makes four hydrogen bond contacts with the $\gamma^{398-411}$ segment main chain (Fig.4). Residues 139-143 ($\gamma^{406-410}$) make up the second wide turn with $Ala^{141}$ ($\gamma^{408}$) separating the two halves.

Because the fibrinogen $\gamma^{398-411}$ segment does not have a hydrophobic core that an independently folding polypeptide segment would have, it is reasonable to ask whether the conformation that is observed is determined solely by interactions with the lysozyme to which it is linked or, possibly, by crystal packing interactions. The simplified representation of the structure of the fibrinogen γ chain segment presented in Fig. 4 shows that there are five main chain hydrogen bond interactions that occur within the $\gamma^{398-411}$ segment and only two main chain hydrogen bond interactions that occur

between the $\gamma^{398-411}$ segment and the lysozyme molecule to which it is attached (labeled A in Fig. 4). Also, there is a close interaction between the side chain of $Asp^{19}$ of lysozyme and the carbonyl group of $Gln^{140}$ ($\gamma^{407}$) (also labeled A in Fig. 4). This would be an unfavorable interaction unless the carboxyl group of $Asp^{19}$ were protonated. There are six hydrogen bond interactions between the $\gamma^{398-411}$ segment and symmetry related chimeric molecules in the crystal lattice. Four hydrogen bond interactions occur between symmetry related $\gamma^{398-411}$ segments (labeled B in Fig. 4) and two hydrogen bond interactions occur between the $\gamma^{398-411}$ segment and symmetry related lysozyme molecules (labeled C and D in Fig. 4). However, the observed hydrogen bond interactions that occur between the side chain of $Gln^{132}$ ($\gamma^{399}$) and the $\gamma^{398-411}$ segment main chain suggest that this structure would exist in the absence of the lysozyme carrier protein or crystal lattice interactions.

In summary, using the carrier protein driven crystallization approach, the structure of receptor recognition segment of the fibrinogen γ chain at a resolution of 2.4 Å was determined.

**Structure of the Fibrinogen Carboxyl Terminal Segment.** As expected, the lysozyme structure was relatively unaffected by the addition of the fibrinogen γ chain segment to its carboxyl terminus. However, the conformation of the fibrinogen $\gamma^{398-411}$ segment could potentially be affected by interactions with the lysozyme to which it is covalently attached. There are five intramolecular hydrogen bond interactions that occur within the $\gamma^{398-411}$ segment and only three hydrogen bond interactions between the $\gamma^{398-411}$ segment and the lysozyme to which it is attached. Since there are more hydrogen bonds between atoms within the fibrinogen $\gamma^{398-411}$ segment than there are with the carrier lysozyme molecule, it was unlikely that the conformation of the $\gamma^{398-411}$ segment was strongly affected by the presence of the lysozyme.

Crystal packing interactions could also potentially affect the conformation of the fibrinogen γ chain segment. The $\gamma^{398-411}$ segment does make six crystal packing contacts. However, three of the six crystal packing contacts are made by the side chain of

$Gln^{140}$ ($\gamma^{407}$). Because this is a relatively long and flexible side chain, it was unlikely that these interactions strongly influence the conformation of the γ chain segment itself. Nevertheless, the potential influence of crystal packing interactions on the structure of the $\gamma^{398-411}$ segment can only be assessed by comparison to the structure of this segment in the context of different crystal packing.

The three-dimensional structure of the carboxyl terminal segment of the human fibrinogen γ chain has a number of hitherto unrecognized features. First, it is organized into a turn that is distinct from the typical β turn suggested by NMR studies of the $\gamma^{400-411}$ peptide in solution (9) or postulated to encompass the RGD cell attachment site of fibronectin (30). Second, this structure does not seem to be stabilized by a salt bridge formed between the ε amino group of $Lys^{139}$ ($\gamma^{406}$) and the carboxyl group of $Asp^{143}$ ($\gamma^{410}$) or $Val^{144}$ ($\gamma^{411}$) as postulated previously (3). The uniqueness of this structure is exemplified by its selective interaction with platelet integrin $\alpha_{IIb}\beta_3$, whereas a multitude of other integrin receptors remain unengaged by this ligand (2). The only other receptor that interacts with this segment of the human fibrinogen γ chain is the staphylococcal clumping factor (7,31). The apposition of acceptor and donor sites for enzymatic crosslinking by Factor XIIIa provides another structural feature for biologic function of the carboxyl terminal segment of the γ chain characterized in this study. The three-dimensional structure presented herein offers powerful information for the development of models of fibrinogen-$\alpha_{IIb}\beta_3$ and fibrinogen-Factor XIIIa interaction and for the design of new inhibitors of these important functions of fibrinogen. In addition, success in using CEW lysozyme as a carrier protein to drive the crystallization of the carboxyl terminal fibrinogen γ chain segment demonstrated the utility of this approach for determining the three-dimensional structure of functional segments of other proteins that are, like fibrinogen, difficult to crystallize.

Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

## REFERENCES

1. Doolittle, R. F. (1984) *Annu. Rev. Biochem.* **53,** 195-229.

2. Hawiger, J. (1994) in *Hemostasis and Thrombosis. Basic Principles and Clinical Practice,* 3D Edition, eds. Colman, R. W., Hirsh, J., Marder, V. J. & Salzman, E. W. (J. B. Lippincot, Philadelphia, PA), pp. 762-796.

3. Kloczewiak, M., Timmons, S., Lukas, T. J. & Hawiger, J. (1984) *Biochemistry* **23,** 1767-1774.

4. Kloczewiak, M., Timmons, S., Bednarek, M. A., Sakon, M. & Hawiger, J. (1989) *Biochemistry* **28,** 2915-2919.

5. Peerschke, E. I. B., Francis, C. W. & Marder, V. J. (1986) *Blood* **67,** 385-390.

6. Farrell, D H., Thiagarajan, P., Chung, D. W. & Davie, E. W. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10729-10732.

7. Strong, D. D., Laudano, A. P., Hawiger, J. & Doolittle, R. F. (1982) *Biochemistry* **21,** 1414-1420.

8. Chen, R. & Doolittle, R. F. (1971) *Biochemistry* **10,** 4486-4491.

9. Blumenstein, M., Matsueda, G. R., Timmons, S. & Hawiger, J. (1992) *Biochemistry* **31,** 10692-10698.

10. Rao, S. P. S., Poojary, M. D., Elliott, B. W., Melanson, L. A., Oriel, B. & Cohen, C. (1991) J. Mol. Biol. 222, 89-98.

11. Jung, A., Sippel, A. E., Grez, M. & Schutz, G. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5759-5763.

12. Miki, T., Yasukochi, T., Nagatani, H., Furuno, M., Orita, T., Yamada, H., Imoto, T. & Horiuchi, T. (1987) *Protein Engineering* **1,** 327-332.

13. Imoto, T., Yamada, H., Yasukochi, T., Yamada, E., Ito, Y., Ueda, T., Nagatani, H., Miki, T. & Horiuchi, T. (1987) *Protein Engineering* **1,** 333-338.

14. Zoller, M. & Smith, M. (1983) *Meth. Enzymol.* **100,** 468-500.

15. Chung, D. W., Chan, W-Y. & Davie, E. W. (1983) *Biochemistry* **22,** 3250-3256.

16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463-5467.

17. Saxena, V. P. & Wetlaufer, D. B. (1970) *Biochemistry* **9,** 5015-5023.

18. Rossmann, M. G. & Blow, D. M. (1962) *Acta Cryst.* **15,** 24-31.

19. Kundrot, C. E. & Richards F. M. (1988) *J. Mol. Biol.* **200,** 401-410.

20. Brunger, A. T. (1993) X-PLOR 3.1, A System for X-ray Crystallography and NMR.

21. Brunger, A. T., Krukowski, A. & Erickson, J. (1990) *Acta Cryst.* **A46,** 585-593.

22. Artymiuk, P. J., Blake, C. C. F., Rice, D. W. & Wilson, K. S. (1982) *Acta Cryst.* **38,** 778-783.

23. Jones, T. A. & Kjeldgaard, M. (1993) O 5.9, A Program for Graphics and Model Building.

24. Brunger, A. T. (1992) *Nature (London)* **355,** 472-474.

25. Hodel, A., Kim, S. H. & Brunger, A. T. (1992) *Acta Cryst.* **A48,** 851-858.

26. Ramachandran, G. N. & Sasisekharan, V. (1968) *Adv. Prot. Chem.* **28,** 283-437.

27. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Cryst.* **26,** 283-291.

28. Yamada, T., Matsushima, M., Inaka, K., Ohkubo, T., Uyeda, A., Maeda, T., Titani, K., Sekiguchi, K. & Kikuchi, M. (1993) *J. Biol. Chem.* **268,** 10588-10592.

29. Davies, D. R., Padlan, E. A. & Sheriff, S. (1990) *Annu. Rev. Biochem.* **59,** 439-473.

30. Ruoslahti, E. & Pierschbacher, M. D. (1987) *Science* **238,** 491-497.

31. McDevitt, D., Francois, P., Vaudraux, P. & Foster, T. J. (1994) *Mol. Microbiol.* **11,** 237-248.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: VANDERBILT UNIVERSITY

(ii) TITLE OF INVENTION: Carrier Protein Driven Crystallization of a Peptide or Polypeptide

(iii) NUMBER OF SEQUENCES: 3

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Needle & Rosenberg, P.C.
(B) STREET: 127 Peachtree Street, N.E., Suite 1200
(C) CITY: Atlanta
(D) STATE: Georgia
(E) COUNTRY: USA
(F) ZIP: 30303

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: PERRYMAN, David G.
(B) REGISTRATION NUMBER: 33,438
(C) REFERENCE/DOCKET NUMBER: 2200.027

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (404) 688-0770
(B) TELEFAX: (404) 688-9880

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gln Gln His His Leu Gly Gly Ala Lys Gln Ala Gly Asp Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (other nucleic acid)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCAACACCAC CTAGGGGGAG CCAAACAGGC TGGAGACGTT TA 42

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (other nucleic acid)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTTAAACG TCTCCAGCCT GTTTGGCTCC CCCTAGGTGG TGTTGCTGCA 50

What is claimed is:

1. A method of crystallizing a peptide or polypeptide comprising linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein and crystallizing the chimeric protein.

2. The method of Claim 1, wherein the peptide or polypeptide is the peptide set forth in SEQ ID NO: 1.

3. The method of Claim 2, wherein the carrier protein is chicken egg white lysozyme.

4. The method of Claim 1, wherein the carrier protein is at least 10,000 $M_r$.

5. The method of Claim 1, wherein the carrier protein is chicken egg white lysozyme, human lysozyme, thioredoxin, glutathione-S-transferase, maltose binding protein, Fab antibody fragment, glycogen phosphorylase, purine nucleotide phosphorylase, beta galactosidase, or a derivative thereof .

6. The method of Claim 5, wherein the carrier protein is chicken egg white lysozyme.

7. The method of Claim 1, wherein the peptide or polypeptide is at least 3 amino acid residues.

8. A method for determining the three dimensional structure of a peptide or polypeptide to a resolution of 1 nanometer or better comprising the steps of linking the peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, crystallizing the chimeric protein, and then analyzing the chimeric protein to determine the three dimensional structure of the crystallized peptide or polypeptide.

9. The method of Claim 8, wherein the analyzing is by x-ray diffraction.

10. The method of Claim 8, wherein the peptide or polypeptide is the peptide set forth in SEQ ID NO: 1.

11. A crystallized chimeric protein, wherein the chimeric protein comprises a peptide or polypeptide linked to a terminus of a crystallizable carrier protein, wherein the peptide or polypeptide is crystallized such that, the three dimensional structure of the crystallized peptide or polypeptide can be determined to a resolution of 1 nanometer or better, and wherein the three dimensional structure of the peptide or polypeptide alone can not be determined to a resolution of 1 nanometer or better.

12. The crystallized chimeric protein of Claim 11, wherein the peptide or polypeptide is the peptide set forth in SEQ ID NO: 1.

13. The crystallized chimeric protein of Claim 11, wherein the carrier protein is at least 10,000 $M_r$.

14. The crystallized chimeric protein of Claim 11, wherein the carrier protein is chicken egg white lysozyme, human lysozyme, thioredoxin, glutathione-S-transferase, maltose binding protein, Fab antibody fragment, glycogen phosphorylase, purine nucleotide phosphorylase, beta galactosidase, or a derivative thereof .

15. The crystallized chimeric protein of Claim 11, wherein the carrier protein is chicken egg white lysozyme.

16. The crystallized chimeric protein of Claim 11, wherein the peptide or polypeptide is at least 3 amino acid residues.

17. The crystallized chimeric protein of Claim 11, wherein the three dimensional structure of the crystallized peptide or polypeptide can be determined to a resolution of 0.24 nanometer or better, and wherein the three dimensional structure of the peptide or polypeptide alone can not be determined to a resolution of 0.24 nanometer or better.

18. A method for designing a first peptide or polypeptide for screening for improved binding to a molecule, comprising the steps of

(1) evaluating the three dimensional structure of a second peptide or polypeptide, which can be the same or different than the first peptide or polypeptide, which was produced by:

a) linking the second peptide or polypeptide to a terminus of a crystallizable carrier protein to form a chimeric protein, and

b) crystallizing the chimeric protein; and

(2) synthesizing a first peptide or polypeptide based on the crystal structure of the second peptide or polypeptide, wherein the first peptide or polypeptide can be screened for having improved binding to the molecule.

19. The method of Claim 18, wherein the peptide or polypeptide is the peptide set forth in SEQ ID NO: 1.

20. The method of Claim 18, wherein the molecule is a receptor.

21. The method of Claim 18, wherein the molecule is Factor XIIIa.

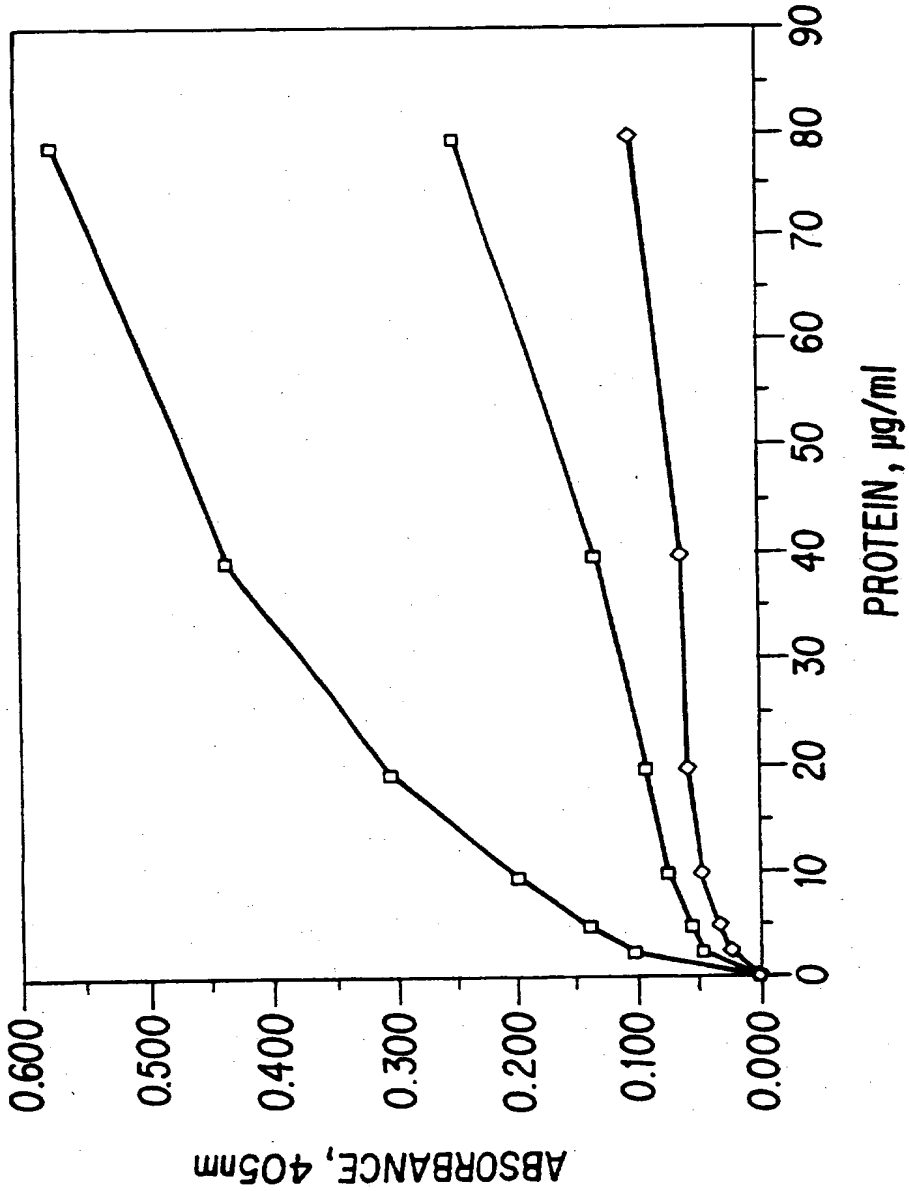
ABSORBANCE, 405nm
0.600
0.500
0.400
0.300
0.200
0.100
0.000
PROTEIN, µg/ml
0
10
20
30
40
50
60
70
80
90

FIG. 1


FIG.2

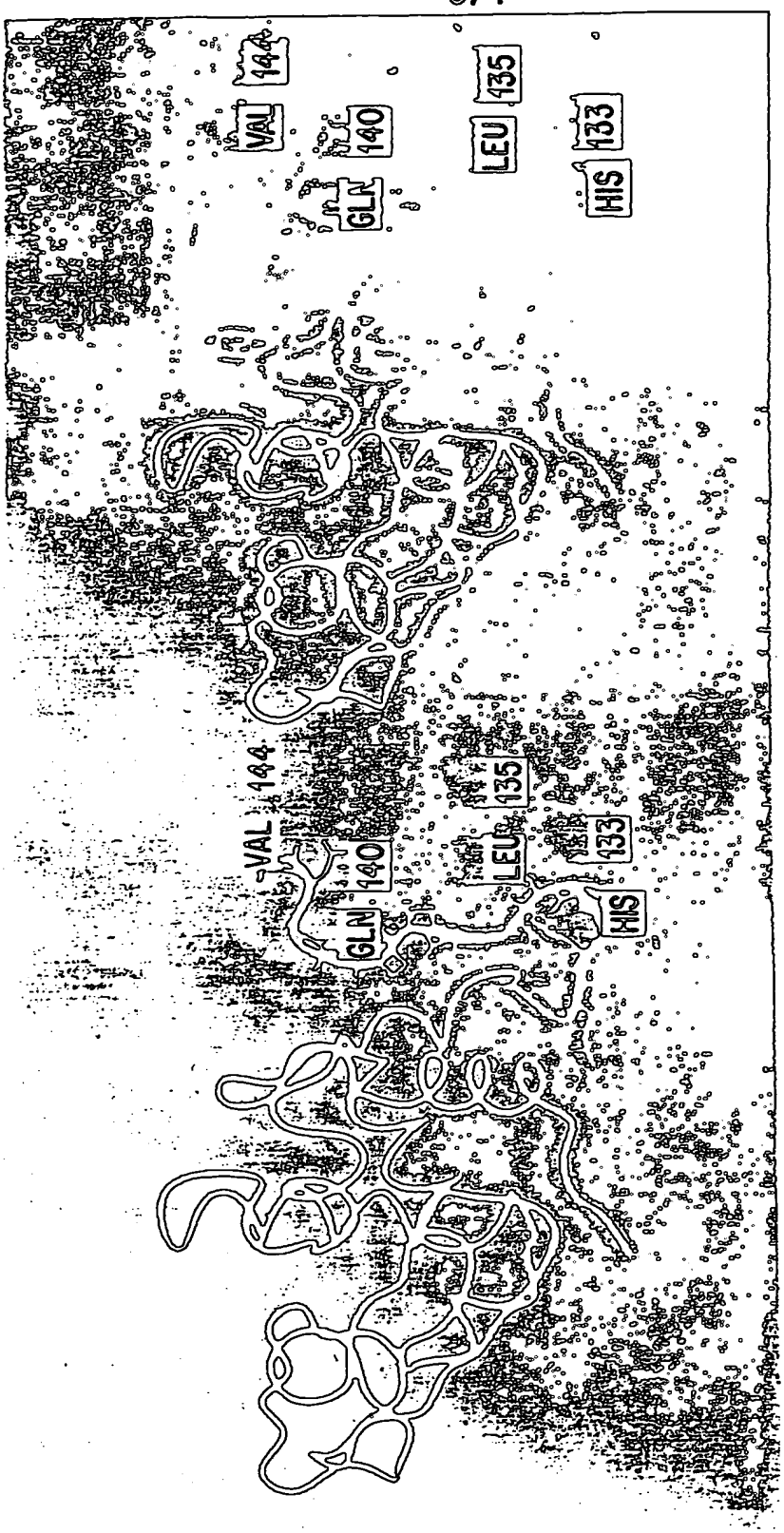


FIG. 3

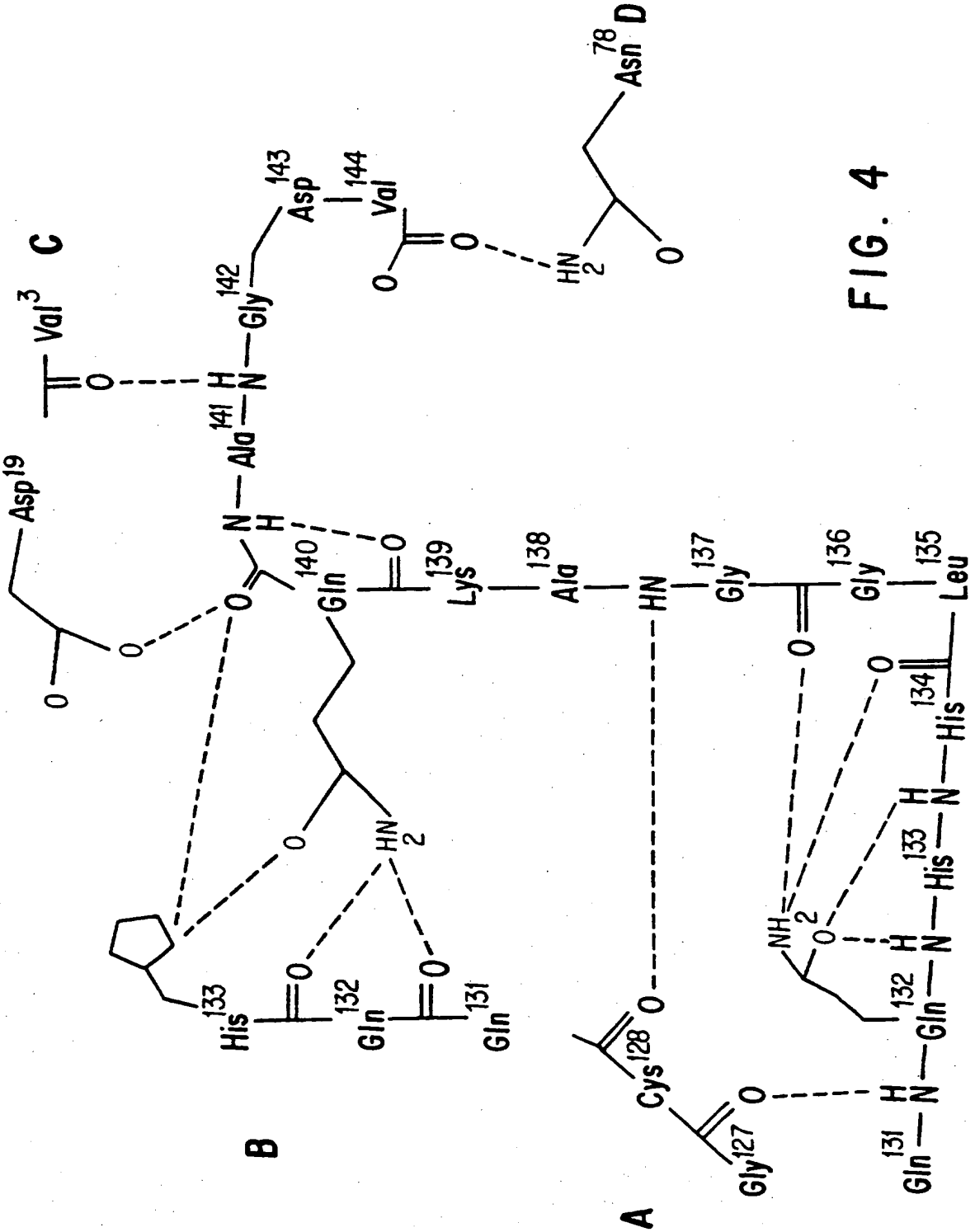
A
B
C
D
Val 3
Asp 19
Ala 141
Gly 142
Asp 143
Val 144
Asn 78
Cys 128
Gly 127
Gln 131
Gln 132
His 133
Gln 131
Lys 139
Gln 140
Ala 138
Gly 137
Gly 136
Leu 135
His 134
His 133
Gln 132

FIG. 4

# INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US95/15714

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C12N 9/36; C30B 19/00, 29/58; C07K 1/04, 17/02
US CL :435/206; 530/350; 117/54, 927

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/206; 530/350; 117/54, 927

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, BIOSIS, CAS, MEDLIN, EMBASE, WPIDS
Search Terms: structure, crystal?, protein#, polypeptide#, bond?, covalent?, attach?

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X<br>----<br>Y | WO, A, 91/02982 (GESELLSCHAFT FUR BIOTECHNOLOGISCHE FORSCHUNG MBH) 07 MARCH 1991, see entire document (English translation provided). | 1, 4, 7-9, 11, 13, 16-17<br>---------<br>2-3, 5-6, 10, 12, 14-15 |
| Y | J. BIOL. CHEM., Volume 268, Number 14, issued 15 May 1993, T. Yamada et al., "Structural and Functional Analyses of the Arg-Gly-Asp Sequence Introduced into Human Lysozyme", pages 10588-10592, see entire document. | 2-3, 5-6, 10, 12, 14-15 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

| | | | |
|---|---|---|---|
| * | Special categories of cited documents: | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| "A" | document defining the general state of the art which is not considered to be of particular relevance | "X" | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| "E" | earlier document published on or after the international filing date | "Y" | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "L" | document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "&" | document member of the same patent family |
| "O" | document referring to an oral disclosure, use, exhibition or other means | | |
| "P" | document published prior to the international filing date but later than the priority date claimed | | |

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 19 MARCH 1996 | 29 MAR 1996 |
| Name and mailing address of the ISA/US<br>Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231<br>Facsimile No. (703) 305-3230 | Authorized officer<br>JON P. WEBER, PH.D.<br>Telephone No. (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15714

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | BIOCHEMISTRY, Volume 28, Number 7, issued 1989, S. Timmons et al., "Antiplatelet "Hybrid" Peptides Analogous to Receptor Recognition Domains on $\gamma$ and $\alpha$ Chains of Human Fibrinogen", pages 2919-2923, see entire document. | 2-3, 5-6, 10, 12, 14-15 |
| A | J. MED. CHEM., Volume 34, Number 7, issued July 1991, K. Appelt et al., "Design of Enzyme Inhibitors Using Iterative Protein Crystallographic Analysis", pages 1925-1934. | 1-17 |
| A | TRENDS BIOCHEM. SCI., Volume 14, issued July 1989, D. Eisenberg et al., "Protein Crystallography: More Surprises Ahead", pages 260-264. | 1-17 |
| Y | J. DRENTH, "PRINCIPLES OF PROTEIN X-RAY CRYSTALLOGRAPHY", published 1994, by Springer-Verlag (N.Y.), pages 7-10, see entire document. | 2-3, 5-6, 10, 12, 14-15 |

# INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US95/15714

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-17

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15714

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-17, drawn to a method of crystallizing a peptide as a chimera with a crystallizable protein, determining its structure by x-ray crystallography, and the crystalline chimeric protein/peptide per se.

Group II, claims 18-21, drawn to a method of designing a second peptide by examining the structure of a first peptide and synthesizing a second peptide.

The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Inventions Groups I and II are drawn to different methods, Group I to crystallizing a protein/peptide chimera and determining its structure, and Group II which is drawn to designing a second peptide based on the structure of a first peptide and synthesizing the second peptide.

It is not necessary to form a chimera and crystallize the chimera to obtain the structure of a peptide which can be used to design a similar structure to the peptide. The peptide could be directly crystallized, or the peptide structure could be determined in solution with high resolution 2D NMR or by molecular modeling techniques.

Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

**IMAGES ARE BEST AVAILABLE COPY.**
**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**